


Assessing Children at Risk: Psychometric Properties of the Motor Behavior Checklist

Journal of Attention Disorders
XX(X) 1–10
© 2013 SAGE Publications
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1087054713484798
jad.sagepub.com


Maria Efstratopoulou¹, Rianne Janssen¹ and Johan Simons¹

Abstract

Objective: The psychometric properties of the Motor Behavior Checklist (MBC) were examined. **Method:** A normative database of primary students ($N = 841$), rated by their physical educators in school settings, were used to investigate the construct validity of the MBC, the internal consistency, the reproducibility, and the interrater agreement. **Results:** Psychometric results supported the model. Confirmatory factor analyses revealed a second-order model with two (Externalizing and Internalizing) broadband domains and seven problems scales: Lack of Attention (10 items), Hyperactivity/Impulsivity (14 items), Rules Breaking (7 items), Low Energy (4 items), Stereotyped Behaviors (2 items), Lack of Social Interaction (10 items), and Lack of Self-Regulation (12 items). **Conclusion:** Although MBC for children is not a diagnostic tool itself, it can provide valid complementary information on attentional, emotional, and developmental problems in children when used by physical educators in school settings. (*J. of Att. Dis.* 2013; XX(X) 1–XX)

Keywords

motor-related behaviors, checklist, physical education, cluster analysis, confirmatory factor analysis, reliability

Introduction

The Importance of Early Diagnosis

Education research indicates that early identification for emotional and/or behavioral problems can help to minimize the long-term harm of mental health disorders and reduce the overall health care burden and costs (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). Detection efforts are particularly critical during the early educational years, when students are most amenable to change in behavioral, social, and academic arenas and before students at risk of emotional and behavioral disorders (EBD) and children with autistic spectrum disorder (ASD), experience negative outcomes within and beyond the school setting (Landrum, Tankersley, & Kauffman, 2003; Lane, 2003; Volkmar, Lord, Bailey, Schultz, & Klin, 2004; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005). Flanagan, Bierman, and Kam (2003) examined the prevalence of child problem profiles at school entry based on behavior problems (e.g., aggression, hyperactivity-inattention, and prosocial skill deficits) and investigated the predictive utility of screening. Findings illustrated that educators who observe different aspects of children's behavior during their lessons were able to identify young children at high risk of school adjustment problems related to attention, conduct, learning, and mood with a great deal of accuracy. In addition, the results indicated interdependency between aggressive, hyperactive-inattentive, and low levels of prosocial behaviors and the existence of

these behaviors made unique contributions to the prediction of later school difficulties in behavioral, academic, and social adjustment domains.

Given the costs associated with EBD, to students themselves, their families, and society as a whole, it is not surprising that reducing the incidence of EBD through systematic screening and comprehensive intervention efforts is a growing area of interest to educational research (Kauffman & Landrum, 2009; Lane, 2007; Nelson, Babyak, Gonzalez, & Benner, 2003).

The Role of Physical Educators in the Assessing Procedure

Results on observational studies suggest that explicit behavioral symptoms can be systematically observed during standardized play procedures (Mol Lous, Wit, De Bruyn, & Riksen-Walraven, 2002). Physical education (PE) lessons and group play situations provide a unique opportunity to observe a child moving, interacting with his or her peers, cooperating or just being on his or her own. PE teachers spend a lot of time with the children, have the flexibility to

¹Catholic University of Leuven, Belgium

Corresponding Author:

Johan Simons, Catholic University of Leuven, Tervuursevest 101, 3001 Heverlee, Belgium.
Email: johan.Simons@faber.kuleuven.be

work with them, and observe their behaviors in several ways (e.g., structured lessons or free play situations) and in several different settings (inside or outside the classroom, at the playground or at the school-yard). Given the fact that evidence for the presence of externalizing and/or internalizing symptoms can be obtained in multiple active situations, and a number of behavioral symptoms can be observed during PE classes and team games (Kashani, Allan, Beck, Bledsoe, & Reid, 1997), PE teachers have the skills and the opportunity to distinguish between maladaptive and general age-related motor-related behaviors among their students.

PE teachers have the knowledge and the skills to focus on the “warning signs” of atypical motor behaviors providing useful information about the development of school-age children. However, there are only a few instruments that use the physical educators as main source of information about children’s development and the majority of them are focusing on movement and motor coordination problems like the Bruininks–Oseretsky Test of Motor Proficiency (BOT-2; Bruininks & Bruininks, 2005), the Test of Gross Motor Development (TGMD; Ulrich, 2000), or the Movement Assessment Battery for Children (MABC-II; Henderson, & Sugden, 2007), which assess gross and fine motor skills, balance, levels of motor skill development as part of psychological test batteries, for making decisions about educational placement, developing and evaluating intervention programs. In addition, none of the existing instruments for physical educators assess a wide array of children’s problematic behaviors, as most of them are focusing only on one specific disorder (e.g., anxiety) which is highly connected with performance in sports like the Physical Education State Anxiety Scale (PESAS; Barkoukis, Tsorbatzoudis, & Grouios, 2008) or are focusing on class management in school settings like the Physical Education Classroom Instrument (PECI; Kulinna, Cothran, & Regualos, 2003).

The Current Study

Considering the need for an instrument that is practical for wide-scale school use by physical educators, assesses a wide array of children’s behaviors, and possesses supportive psychometric evidence, this study aimed to examine the psychometric properties of the Motor Behavior Checklist (MBC) in a elementary school-aged sample of children, namely, the factorial validity, the internal consistency, the test–retest reliability, and the interrater reliability.

Method

The Preliminary List

Preliminary list included PE teachers’ reports about children’s problematic behaviors in school settings, and observable motor-related behaviors from the official psychiatric

criteria about children’s psychopathology (*Diagnostic and Statistical Manual of Mental Disorders [DSM-IV-TR]*; American Psychiatric Association, 2000; International Statistical Classification of Diseases and Related Health Problems–Tenth Revision [ICD-10]; World Health Organization, 1992). During instrument development, every attempt was made to be sensitive to the varied contexts and children who participate in PE classes and some of the items represented behaviors unique to PE settings. A preliminary version of the MBC for children containing 85 items and a 5-point response scale format was used ranging from 0 (*never*) to 4 (*almost always*) for each of the items to give the rater the possibility to provide information about the frequency of the exhibited behavior.

Participants

The data analyzed were collected from a randomly selected sample ($N = 841$) of elementary school-age children. School review broad approval was obtained as well as appropriate consent/assent from participants and their parents. The data collected were anonymous and only codes about demographic characteristics of the participants were used. Overall sample consisted of 421 (50.1%) girls and 420 (49.9%) boys, ranging in age from 6 to 11 years ($M = 8.4$ years, $SD = 1.7$ years) and they were of Greek nationality (99%). The data derived from 35 typical Greek elementary schools widely spread across Greece selected so that the sample distribution would be representative of the urban and rural population. The schools were located in urban areas (63.3%) and in rural areas and islands (36.7%). The PE teachers ($n = 62$) of the schools who participated in this study were 35 females (56%) and 27 males (44%), with mean age 39.4 years ($SD = 6.2$ years) and mean teaching experience 7.2 years ($SD = 3.4$ years). The participants were asked to randomly select four children (2 boys and 2 girls) from each class and rate them using the 85-item preliminary version of MBC for children. The initial data were randomly divided into two sample groups. Sample 1 ($n = 426$) was used for examining the structure validity of the list and Sample 2 ($n = 415$) was used to cross validate the data and further assess the model fit. An overview of children’s characteristics from confirmatory factor analyses (CFAs) and reliability studies are presented in Table 1.

Data Analysis

The aims of study were twofold: (a) investigation of the factorial structure via confirmatory analyses and (b) examination of the reliability of the list. More specifically, the investigation of the structure validity was conducted into three different methodological steps: (a) initial examination of the factor structure of the list, (b) selection of the items based on specific criteria, and (c) investigation of the

Table 1. Children's Characteristics Participating per Study.

Variables	CFA factor structure (<i>n</i> = 426)	CFA assessing model fit (<i>n</i> = 415)	Test–retest reliability (<i>n</i> = 129)	Interrater reliability (<i>n</i> = 126)
Gender				
Boys	(50%)	(50.1%)	(52%)	(55%)
Girls	(50%)	(49.9%)	(47%)	(45%)
Nationality				
Greek	(99.3%)	(99.4%)	(86%)	(93%)
Age in years ^a	8.7 (1.7)	8.7 (1.7)	9.7 (1.2)	7.6 (1.5)

Note: CFA = confirmatory factor analysis.

^aStandard deviations are in parentheses.

adequacy of the model fit using cross-validation. Concerning the reliability of the list, the internal consistency, the reproducibility, and the interrater agreement were also examined.

Factor structure. The initial data were divided randomly into two data samples. Data from Sample 1 (*n* = 426) were used to run a CFA using maximum-likelihood method (LISREL 8; Jöreskog & Sörbom, 1993). Modifications to the hypothesized factor model were made based on the correlations among factors.

Selection of the items. Taking under consideration the selection criteria proposed by Marsh, Ellis, Parada, Richards, and Heubeck (2005), the standardized loadings, the modifications indices, and the item total correlations were examined to produce a more concise instrument within a parsimonious model. A confirmatory analysis was performed to examine a second-order factor model with seven factors using the 85 items of the list. In case of inadequate fit for the hypothesized model, modifications to the CFA were performed based on the analyses of items. The modifications made were based on high factor loadings, correlated uniqueness within each factor and inter correlations between items within each factor. More specifically, items that (a) best measured each factor having high standardized factor loadings ($\geq .50$) and (b) had minimal cross loadings on other factors as assessed via modification indices were chosen. Caution was taken not to reduce the number of subscales or the number of items within each scale so drastically that it led to construct under presentation that could mask the intended purpose and validity of the measure (Messick, 1995). The modifications made consisted of freeing up cross loadings and correlated uniqueness within each factor until a reasonable fit was obtained.

Assessing model fit. To confirm the adequacy of the model which revealed via the CFAs, we performed a cross-validation procedure. Data from Sample 2 (*n* = 415) were

used to run a second-order factor confirmatory analysis using maximum-likelihood (LISREL 8; Jöreskog & Sörbom, 1993). Because the chi-square statistic frequently yields false positives when large samples are analyzed (Floyd & Widaman, 1995), the comparative fit index (CFI), the normed fit index (NFI), the goodness-of-fit index (GFI), the root mean of approximation (RMSEA), and the standardized root mean square residual (SRMS) indexes were used to evaluate the fit of the data.

Internal consistency. Internal consistency was estimated to measure the extent to which items in a subscale are correlated (homogeneous), thus measuring the same concept. Cronbach's alpha coefficients were calculated of the subscales using the initial data of children (*N* = 841). Alpha values estimated separately for the Externalizing and Internalizing scales and the seven problem subscales.

Reproducibility. To determine test–retest reliability, intra-class correlation coefficients (ICC) were calculated using a sample of 129 elementary school children, 61 girls (47%), and 68 boys (52%) who were rated twice by their physical educators at schools. The children had mean age 8.51 years (*SD* = 1.75), 111 (86%) had the Greek nationality, and they were attending nine typical elementary schools from Athens and Thessaloniki. The participant physical educators, 7 females and 4 males, recorded their student's motor-related behaviors during PE lessons in school environment using the MBC for children twice within 1 month.

Interrater reliability. Data from 22 physical educators (14 females and 8 males) from 11 elementary schools from Athens and Thessaloniki, who rated 126 of their students using the MBC, were used to assess interrater reliability of the checklist. The participant children were 67 boys (53%) and 59 girls (47%), with mean age 8.51 years (*SD* = 1.75) and 118 (92%) were of Greek nationality. From each school, two physical educators working independently with the same students but in different settings were involved. One

was the traditional PE teacher working only in movement situations, involving sports, and the second one was working in both settings, inside the classroom, giving theoretical information about regulations in sports and nutrition, and outside the classroom working in team games.

Results

Factor Structure

Examination of the nine-factor model revealed high correlation between factor Disobedience and factor Aggressiveness ($r = .98$) and between factor Hyperactivity and factor Impulsivity ($r = .95$). It was assumed that these factors were so similar in content describing and assessing aspects of the same construct and therefore had to be reduced into two. Factor Disobedience (8 items) and factor Aggressiveness (7 items) were merged into one factor containing the sum of their items ($n = 15$ items) and it was named "Rules Breaking" as the most of the items were describing aggressive behaviors mainly connected with disobedience and violation of rules in school environment. Factor Hyperactivity (6 items) and factor Impulsivity (9 items) were also merged into one factor containing the sum of their items ($n = 15$ items) and named "Hyperactivity/Impulsivity" containing items describing hyperactive and impulsive behaviors. The other five factors were Low Energy ($n = 4$ items), with items describing decreased activity; Stereotyped Behaviors ($n = 6$ items), with items describing repeated patterns of activity; Lack of Attention ($n = 10$ items), containing items describing problems in attention and lack of concentration; Lack of Social Interaction ($n = 16$ items), containing items describing problems in communication and social interaction with teachers and peers; and Lack of Self-Regulation ($n = 19$ items), containing mainly items describing anxiety and inability of the child to regulate behavior.

Selection of the Items

A second-order CFA indicated the existence of two higher orders Externalizing and Internalizing factors containing the seven problems subscales. The indices of fit for the initial model were as follows: RMSEA = .092, CFI = .87, NFI = .90, GFI = .89, and SRMS = .068. These indices were considered rather low, and to improve the model fit and reduce the number of items, based on the criteria proposed by Marsh et al. (2005), the following modifications were made: (a) Items with low factor loadings ($\leq .50$) were deleted and (b) items with cross loadings were excluded from the initial list. In this way, a new sorter list containing 59-items was developed. More specifically, the reduced model consisted of factor Rules Breaking (7 items), Hyperactivity/Impulsivity (14 items), Lack of Attention (10 items), Low

Energy (4 items), Stereotyped Behavior (2 items), Lack of Social Interaction (10 items), and Lack of Self-Regulation (12 items). The reduced model presented an optimal level of fit to the data (RMSEA = .074, CFI = .97, NFI = .96, GFI = .93, and SRMS = .55) and was characterized by significant and substantial loadings (i.e., ranging from .57 to .81). The second-order CFA model for the MBC is presented in Figure 1.

Correlation between the two higher factors, Externalizing and Internalizing problems, was $r = .32$. The estimated correlations between the Externalizing factor and Rules Breaking, Hyperactivity/Impulsivity, Lack of Attention factors were $r = .95$, $r = .98$, and $r = .85$, respectively. In addition, the correlations between the Internalizing domain and the four factors—Low Energy, Stereotyped Behavior, Lack of Social Interaction, and Lack of Self-Regulation—were $r = .81$, $r = .80$, $r = .95$, and $r = .97$, respectively.

Model Fit

We performed cross-validation on the second half of the sample to confirm that the reduced model presented an optimal level of fit to the data. Results from the CFA on the second half of the data ($n = 415$) and examination of the fit indices supported further the adequacy of the reduced model fit. More specifically, the CFI was .96, the NFI was .95, and the GFI was .92. In addition, the RMSEA was .072 and the SRMS value was .054. According to Hu and Bentler (1999), values less than .80 for RMSEA and SRMS indices are acceptable and indicate good fit for the factor model. Correlations among factors and indices of fit are presented in Table 2.

Internal consistency. Alpha values for all the subscales were excellent suggesting that the list was homogeneous in content. More specifically, for the factor Rules Breaking (7 items), alpha value was .95; for Low Energy (4 items), alpha = .82; for Stereotyped Behavior (2 items), alpha = .85; for Hyperactivity/Impulsivity (14 items), alpha = .95; for Lack of Attention (10 items), alpha = .95; for Lack of Social Interaction (10 items), alpha = .94; and for Lack of Self-Regulation (12 items), the alpha coefficient was .91. In addition, for the Externalizing scale (31 items), alpha coefficient was .93, and for the Internalizing scale (28 items) the coefficient alpha was .91.

Reproducibility. ICC for each of the seven subscales were calculated separately. For the Rules Breaking scale, the ICC was .87; for the Low Energy factor .78; for the Stereotyped Behavior, it was .82; for factor Hyperactivity/Impulsivity, it was .90; for the Lack of Attention, it was .89; for Lack of Social Interaction, it was .85; and for the Lack of Self-Regulation, the ICC coefficient was .83. For the Externalizing scale, the ICC coefficient was .87, and for the Internalizing

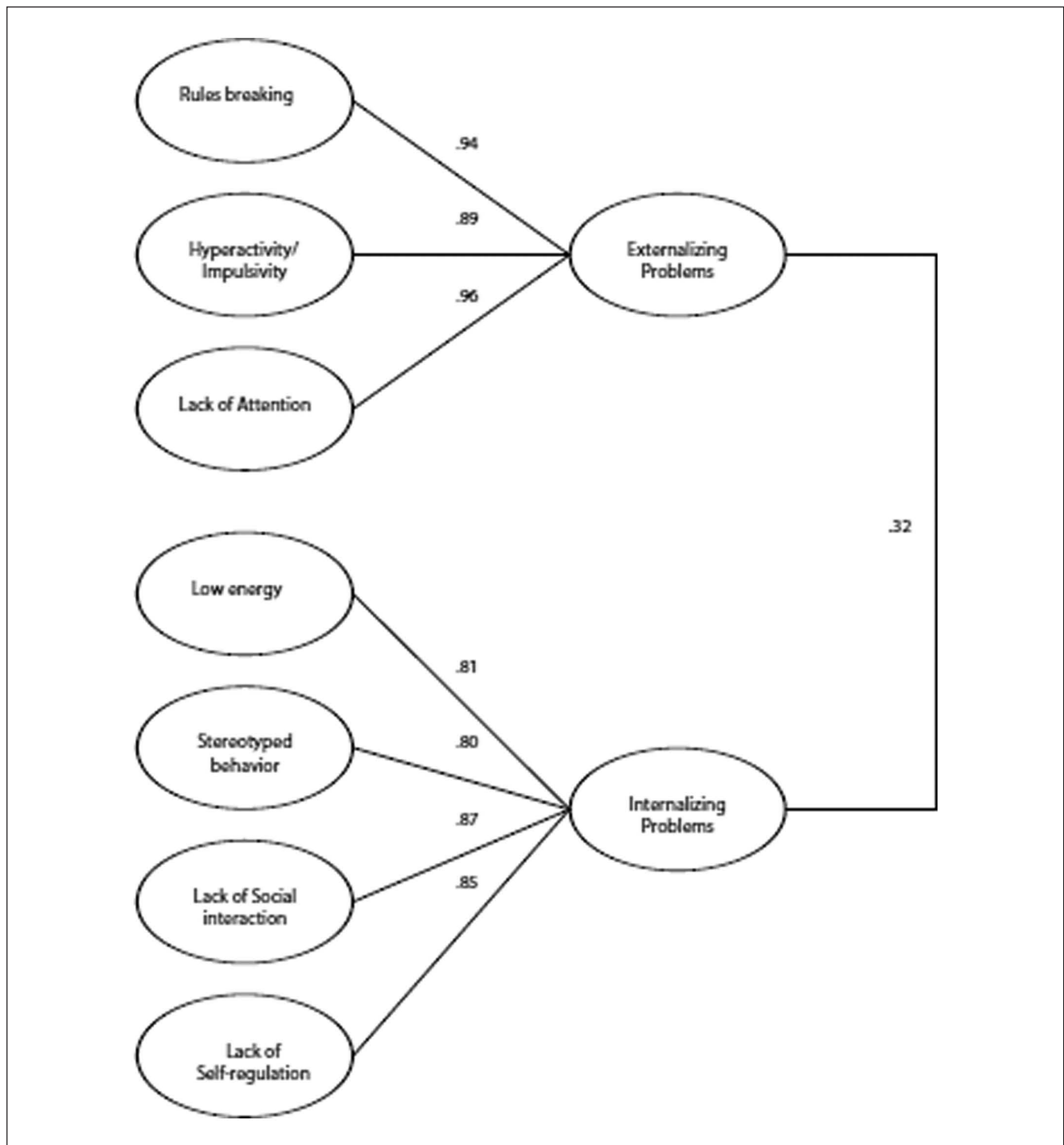


Figure 1. Second-order CFA of the MBC.

Note: CFA = confirmatory factor analysis; MBC = Motor Behavior Checklist.

scale, the ICC coefficient was .81. All correlations were highly significant at $p < .001$ (see Table 3).

Interrater reliability. The ICC coefficients were significant at $p < .001$, ranging from .75 (Low Energy) to .91

(Lack of Attention). More specifically, for the scales of the Externalizing domain, the interrater agreement was for the Rules Breaking factor ICC = .88, for the Lack of Attention factor ICC = .91, and for the Hyperactivity/Impulsivity factor ICC = .88. In addition, lower but statistical signifi-

Table 2. Fit Statistics and Correlations Among Scales of the MBC.

Scales	Rules Breaking	Low Energy	Stereotyped Behaviors	Hyperactivity/Impulsivity	Lack of Attention	Lack of Social Interaction	Lack of Self-Regulation	Externalizing Behaviors	Internalizing Behaviors
Rules Breaking	—								
Low Energy	.15	—							
Stereotyped Behaviors	.03	.59*	—						
Hyperactivity/Impulsivity	.66*	.22	.24	—					
Lack of Attention	.56*	.19	.14	.83*	—				
Lack of Social Interaction	.21	.65*	.53*	.31	.27	—			
Lack of Self-Regulation	.21	.66*	.62*	.32	.27	.92*	—		
Externalizing Behaviors	.94*	.23	.14	.89*	.96*	.34*	.32	—	
Internalizing Behaviors	.22	.81*	.80*	.23	.28	.87*	.85*	.32	—
Indices of fit									
Model	RMSEA	CFI	NFI	GFI	SRMS				
59-items MBC	.05	.97	.96	.93	.07				

Note: MBC = Motor Behavior Checklist; RMSEA = root mean of approximation; CFI = comparative fit index; NFI = normed fit index; GFI = goodness-of-fit index; SRMS = standardized root mean square residual.

* $p < .001$.

Table 3. Reliability Coefficients.

Scales	Items	Internal consistency (Cronbach's α)	Test-retest reliability (ICC)	Interrater reliability (ICC)
		($N = 841$)	($n = 129$)	($n = 126$)
Rules Breaking	7	.95	.87*	.88*
Low Energy	4	.82	.78*	.75*
Stereotyped behaviors	2	.85	.82*	.85*
Hyperactivity/Impulsivity	14	.96	.90*	.88*
Lack of Attention	10	.95	.89*	.91*
Lack of Social Interaction	10	.92	.85*	.74*
Lack of Self-Regulation	12	.91	.83*	.81*
Externalizing items	31	.93	.87*	.78*
Internalizing items	28	.91	.81*	.71*

Note: ICC = intraclass correlation coefficient.

* $p < .001$.

cant correlations were assessed for Internalizing factors. The correlation agreement for the Low Energy factor was $ICC = .75$, for the Stereotyped Behaviors was $ICC = .85$, for the Lack of Social Interaction was $ICC = .74$, and for the Lack of Self-Regulation factor was $ICC = .81$. In addition, ICC coefficient for the Externalizing scale was .78, and for the Internalizing scale, the ICC coefficient was .71. Internal consistency, test-retest reliability, and interrater reliability coefficients are presented in Table 3.

Summary of the Results

Aim of study was to investigate the key psychometric properties, namely, the structure validity, the internal consistency, the temporal stability, and the interrater agreement of a new scale (i.e., the MBC) for the assessment of emotional, behavioral, and developmental disorders in elementary school-age children by the PE teacher. A series of CFAs

revealed a second-order model with two (Externalizing and Internalizing) broadband factors and seven problems scales. The items belonging to the two main factors and the seven problem scales are presented in Table 4.

The internal consistency was high for each scale suggesting that the list was homogeneous in content. In addition, the reproducibility and the interrater agreement were excellent suggesting that the MBC for children is an instrument with high temporal stability and high correlation agreement when used by physical educators in school settings.

General Discussion

A New Instrument for PE Teachers

Students with EBD include a wide range of children and youth, including those with externalizing and internalizing

Table 4. Motor Behavior Checklist for Children: Items per Problem Scale.

Externalizing scales			Internalizing scales			
I. Rules Breaking	II. Hyperactivity Impulsivity	III. Lack in Attention	IV. Low Energy	V. Stereotyped Behavior	VI. Lack of Social Interaction	VII. Lack of Self-Regulation
1. Doesn't comply with the rules especially in a game situation	2. Has difficulty waiting his or her turn to perform	3. Is careless	4. Displays tiredness after even a minimum effort	5. Displays stereotyped body movements, which include the hands (e.g., clapping, finger flicking)	6. Displays impairment in gestures which regulate social interaction	7. Displays persistent preoccupation with parts of objects
8. Displays disobedience for his or her teacher	9. Displays hyperactivity during lesson	10. Has a difficulty in concentration	11. Feels dizzy, unsteady, lightheaded, or faint	12. Displays repetitive patterns of activities	13. Displays avoidance of age-appropriate social activities	14. Shows no interest for the lesson
15. Is aggressive toward leadership figures	16. Interrupts others (e.g., butts into conversation)	23. Appears as if he or she did not hear what has just been said	18. Displays low energy	19. Doesn't show objects he or she finds interesting	20. Displays marked impairment in the use of nonverbal behaviors (eye-to-eye contact)	26. Displays impairment in facial expression
21. Is negative toward his or her classmates	22. Interrupts others (e.g., butts into games)	29. Avoids or has strong dislike for activities that require close concentration	24. Shows decreased activity	25. Doesn't bring objects he or she finds interesting	30. Shrinks from contact with others	31. Does not recognize his or her fear as excessive
27. Blames others for his or her mistakes	28. Doesn't care for equipment	34. Displays difficulty in concentration in the beginning of the lesson		35. Displays lack to communicate with his or her classmates	36. Fears standing in a line	41. Displays anxiety that may expressed by crying, tantrums, freezing or clinging
32. Plays too rough during team games	33. May shift from one uncompleted activity to another	39. Has difficulties organizing tasks		40. Is isolated by his or her classmates	45. Displays difficulty making decisions	49. Finds it difficult to control worry
37. Displays bullying tendency toward his or her classmates	38. Engages in dangerous activities without consideration of possible consequences	43. Makes careless mistakes in activities		44. Attempts to remain close to familiar adults	52. Gets upset when fails to accomplish a task	54. Has difficulties organizing group activities
	42. Appears to be driven by a motor	47. Fails to pay close attention to details		48. Doesn't want physical contact	58. Displays lack of varied make-believe play	
	46. Has difficulty to play or engage quietly in leisure activities	55. Avoids or has strong dislike for activities that require organizational demands		51. Doesn't participate actively in simple social play		
	50. Gets upset when loses					
	53. Overestimates his/her capabilities					
	56. Displays impulsive behavior					
	57. Touches things he or she is not supposed to touch					
	59. Loses his or her temper					

behavior problems (Morris, Shah, & Morris, 2002; Walker, Ramsey, & Gresham, 2004). These students often have broad-based needs because of their behavioral, social, and academic deficits which often do not improve over time (Lane, 2007; Mattison, Hooper, & Glassberg, 2002; Nelson et al., 2003). Because not all students with EBD will necessarily require special education, it is very important that educators and school administrators be prepared to implement systematic screening efforts to identify students who may show early signs of EBD. As such, the first step is to implement systematic screening tools to identify students who might benefit from more focused supports (Lane, 2007).

Physical educators have the advantage of observing the child within a peer group, allowing these experts in movement situations to distinguish between maladaptive and typical age-related behaviors. This study fills an important gap in the literature as physical educators lack a practical and reliable instrument for providing useful and valid information about children's with behavioral and emotional problems on the basis of their motor-related behaviors, despite the fact that a lot of useful information could be obtained through observation during PE lessons in school settings or free play situations (Mol Lous et al., 2002).

MBC for children is a new practical and useful measure for assessing externalizing and/or internalizing problems in elementary school-age children by the PE teacher. From this point of view, the instrument could be used to provide valuable additional information about child's problematic behavior and help physical educators in their important decision to refer or not students for further evaluation by the diagnostic teams. Although the MBC for children is not designed to be used as a diagnostic tool in clinical settings, however, the data provided by the instrument could be useful as complementary information during assessment procedures. Rating the child on a number of motor-related behaviors, a lot of valuable information concerning the global behavioral status of the child could help pediatricians and school psychologists, during their psychological evaluation and especially when psychomotor intervention programs and behavioral interventions are proposed to applied.

Psychometric Properties of the MBC for Children

The aim of this study was the evaluation of the key psychometric properties of the new instrument: structure validity, internal consistency, reproducibility, and interrater agreement. A series of CFAs established a second-order model with two (Externalizing and Internalizing) broadband domains and seven problems scales (Rules Breaking, Hyperactivity/Impulsivity, Lack of Attention, Low Energy, Stereotyped Behaviors, Lack of Social Interaction, and Lack of Self-Regulation). Items selection and items reduction per scale were based on statistical techniques (i.e., factor loadings, correlated uniqueness), and the selected items

reflected areas that are important to the target population that is being studied. Therefore, the target populations (elementary students) as well as the target experts (PE teachers) were involved during item selection.

Results from the internal consistency revealed high and significant values for each problem scale suggesting that the list was homogeneous in content. Results from test-retest study supports evidence that the list is characterized by satisfactory short-term stability. The level of correspondence between the test and retest was significant, but yet, it should be noted that the time lapse between the two administrations was relatively short (2 weeks) and future research should verify whether the MBC results remain stable over longer period of time. In addition, the interrater agreement was significant for all the problems scales, but the higher correlations agreements between the two physical educators were on the externalizing problems scales and especially on Rules Breaking and Hyperactivity/Impulsivity scales, when the lowest agreement were noticed on the Lack of Social Interaction scale and the Low Energy scale. The different educational settings between the two observers could partly justify the lower agreement on these Internalizing problem scales as social interaction and decreased activity are not easily observed in classroom settings.

An issue that deserves to be discussed is the incremental validity of MBC as a new instrument in the Greek culture. The fact that MBC for children was developed based on a theoretical and procedural framework derived from pilot studies in Greece makes it more ecological valid and appropriate in identifying culture specific aspects of a construct (Tsaousis & Georgiadis, 2009). Furthermore, the instrument contains items that were derived from the reports of Greek physical educators and as is demonstrated in this study, the MBC for children has sound psychometric properties.

Limitations and Future Research

Participants were typical Greek elementary school-age children. It must be noted that findings may differ with a more diverse sample of clinical populations from psychiatric centers. In addition, future research studies are needed to investigate the ability of the list to predict children with disorders based on MBC scores on separate problem scales of the instrument. Future research efforts could investigate the degree to which the MBC is a tool sensitive and specific in identifying students with (a) externalizing problems and/or (b) Internalizing problems with accuracy.

Conclusion

Taking into consideration that early identification for emotional, behavioral, and/or developmental problems can help to minimize the long-term harm of mental disorders and reduce the overall health care burden and costs (Aos et al.,

2004), the MBC for children could be used for various educational purposes, including research projects and intervention programs. The findings of this study are quite encouraging for the future use of MBC for children in the Greek population. Psychometric results supported the model suggesting that MBC for children is a new instrument homogeneous in content, with high temporal stability and high interrater agreement that can provide useful and reliable ratings on behavioral and emotional problems in children when used by PE teachers in school settings. In addition, recent research studies (Efstratopoulou, Janssen, & Simons, 2012a; Efstratopoulou, Janssen, & Simons, 2012b) have also established the discriminant and concurrent validity of the instrument.

Implication for practical use of the MBC for children in school settings could investigate children's problematic motor-related behavior and the effectiveness of intervention programs aiming to reduce inappropriate behavior. The information provided by the MBC may contribute to physical educators in developing class management techniques and assess the effectiveness of their educational interventions with a pre-post administration of the instrument. Moreover, one very important issue that is connected with special education settings is that the information provided, when assessing children's deviant behaviors in a valid and systematic way within elementary school settings, may help PE teachers to decide about the referral or not of children for further diagnostic evaluation.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia: Washington State Institute for Public Policy.
- Barkoukis, V., Tsorbatzoudis, H., & Grouios, G. (2008). Construct validity of the Physical Education Anxiety Scale: A multi-trait-multimethod approach. *Perceptual and Motor Skills*, 107, 651-662.
- Bruininks, R., & Bruininks, B. (2005). *Bruininks-Oseretsky Test of Motor Proficiency* (2nd ed.). Minneapolis, MN: NCS Pearson.
- Efstratopoulou, M., Janssen, R., & Simons, J. (2012a). Agreement among physical educators, teachers and parents on children's behaviors: A multitrait-multimethod design approach. *Research in Developmental Disabilities*, 33, 1343-1351. doi:10.1016/j.ridd.2012.03.015
- Efstratopoulou, M., Janssen, R., & Simons, J. (2012b). Differentiating children with attention-deficit/hyperactivity disorder, conduct disorder, learning disabilities and autistic spectrum disorders by means of their motor behavior characteristics. *Research in Developmental Disabilities*, 33, 196-204.
- Flanagan, K. S., Bierman, K. L., & Kam, C. M. (2003). Identifying at-risk children at school entry: The usefulness of multibehavioral problem profiles. *Journal of Clinical Child & Adolescent Psychology*, 32, 314-326.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 6, 141-148.
- Henderson, S. E., & Sugden, D. A. (2007). *Movement assessment battery for children* (2nd ed.). London, England: The Psychological Corporation.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8.11*. Chicago, IL: Scientific Software International.
- Kashani, J. H., Allan, W. D., Beck, N. C., Bledsoe, Y., & Reid, J. C. (1997). Dysthymic disorder in clinically referred preschool children. *Journal of the American Academy of Child and Adolescents Psychiatry*, 36, 1426-1433.
- Kauffman, J. M., & Landrum, T. (2009). *Characteristics of emotional and behavioral disorders of children and youth* (8th ed.). Columbus, OH: Merrill.
- Kulinna, P. H., Cothran, D. J., & Regualos, R. (2003). Development of an instrument to measure student disruptive behavior. *Measurement in Physical Education and Exercise Science*, 7, 25-41.
- Landrum, T. J., Tankersley, M., & Kauffman, J. M. (2003). What is special about special education for students with emotional and behavioral disorders? *Journal of Special Education*, 37, 148-156.
- Lane, K. L. (2003). Identifying young students at risk for antisocial behavior: The utility of "teachers as tests." *Behavioral Disorders*, 28, 360-389.
- Lane, K. L. (2007). Identifying and supporting students at risk for emotional and behavioral disorders within multi-level models: Data driven approaches to conducting secondary interventions with an academic emphasis. *Education & Treatment of Children*, 30, 135-164.
- Marsh, H. W., Ellis, L., Parada, L., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment*, 17, 81-102.
- Mattison, R. E., Hooper, S. R., & Glassberg, L. A. (2002). Three-year course of learning disorders in special education students classified as behavior disordered. *Journal of the American Academy of Child & Adolescent Psychology*, 41, 1454-1461.
- Messick, S. (1995). Validity of psychological measurement: Validation of inferences from a person's responses and performance in scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.

- Mol Lous, A., Wit, C. A. M., De Bruyn, E. J., & Riksen-Walraven, J. M. (2002). Depression markers in young children's play: A comparison between depressed and non depressed 3 to 6 years old in various play situations. *Journal of Child Psychology and Psychiatry*, 43, 1029-1038.
- Morris, R. J., Shah, K., & Morris, Y. P. (2002). Internalizing behavior disorders. In K. Lane, F. Gresham & T. O'Shaughnessy (Eds.), *Interventions for children with or at risk for emotional and behavioral disorders* (pp. 223-241). Boston, MA: Allyn & Bacon.
- Nelson, N. J., Babyak, A., Gonzalez, J., & Benner, G. J. (2003). An investigation of the types of problem behaviors exhibited by K-12 students with emotional or behavioral disorders in public school settings. *Behavioral Disorders*, 28, 348-359.
- Tsaousis, I., & Georgiadis, S. (2009). The Greek Personality Adjective Checklist. *European Journal of Psychological Assessment*, 25, 164-174.
- Ulrich, D. (2000). *Test of gross motor development: Examiner's manual* (2nd ed.). Austin, TX: PRO-ED.
- Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 45, 135-170.
- Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbance receiving special education services. *Journal of Emotional and Behavioral Disorders*, 13, 79-96.
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in school: Evidence-based practices*. Belmont, CA: Wadsworth.
- World Health Organization. (1992). *The ICD-10 classification of mental and behavioral disorders: Diagnostic criteria for research*. Geneva, Switzerland: Author.

Author Biographies

Maria Efstratopoulou, PhD, is a doctor in biomedical sciences, Department of Kinesiology and Rehabilitation Sciences, KULeuven, Belgium.

Rianne Janssen, PhD, is a professor in the Faculty of Psychology and Educational Sciences, Research Unit in Quantitative Psychology and Individual Differences, KULeuven, Belgium.

Johan Simons, PhD, is a professor in the Department of Rehabilitation Sciences, Research Unit in Adapted Physical Activity and Psychomotor Rehabilitation, KULeuven, Belgium.